计算机科学
COMPUTER SCIENCE

SGPot:一种基于强化学习的智能电网蜜罐框架

王毓贞, 宗国笑, 魏强

引用本文

王毓贞, 宗国笑, 魏强. SGPot:一种基于强化学习的智能电网蜜罐框架[J]. 计算机科学, 2024, 51(2): 359-370.

WNAG Yuzhen, ZONG Guoxiao, WEI Qiang. SGPot:A Reinforcement Learning-based Honeypot Framework for Smart Grid [J]. Computer Science, 2024, 51(2): 359-370.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.221100187

SGPot:一种基于强化学习的智能电网蜜罐框架

王毓贞　宗国笑　魏　强
信息工程大学网络空间安全学院　郑州 450001
(wangyuzhen@mail.nwpu.edu.cn)

摘　要　随着工业4.0的快速推进,与之互联的电力数据采集与监视控制(Supervisory Control and Data Acquisition,SCADA)系统逐渐趋于信息化和智能化。由于这些系统本身具有脆弱性以及受到攻击和防御能力的不对等性,使得系统存在各种安全隐患。近年来,针对电力攻击事件频发,亟需提出针对智能电网的攻击缓解方法。蜜罐作为一种高效的欺骗防御方法,能够有效地收集智能电网中的攻击行为。针对现有的智能电网蜜罐中存在的交互深度不足、物理工业过程仿真缺失、扩展性差的问题,设计并实现了一种基于强化学习的智能电网蜜罐框架——SGPot,它能够基于电力行业真实设备中的系统不变量模拟智能变电站控制端,通过电力业务流程的仿真来提升蜜罐欺骗性,诱使攻击者与蜜罐深度交互。为了评估蜜罐框架的性能,搭建了小型智能变电站实验验证环境,同时将SGPot和现有的GridPot以及SHaPe蜜罐同时部署在公网环境中,收集了30天的交互数据。实验结果表明,SGPot收集到的请求数据比GridPot多20%,比SHaPe多75%。SGPot能够诱骗攻击者与蜜罐进行更深度的交互,获取到的交互会话长度大于6的会话数量多于GridPot和SHaPe。

关键词:智能电网;强化学习;智能交互;主动防御;蜜罐
中图分类号　TP393

SGPot:A Reinforcement Learning-based Honeypot Framework for Smart Grid

WNAG Yuzhen,ZONG Guoxiao and WEI Qiang
School of Cyberspace Security,Information Engineering University,Zhengzhou 450001,China

Abstract　With the rapid advancement of Industry 4.0,the supervisory control and data acquisition(SCADA) system,which is interconnected with Industry 4.0,is gradually becoming more informationized and intelligent. There are various security hazards in the SCADA system caused by the vulnerability of the system and the disparity in attack and defense capability. Due to the frequency of power attacks in recent years,there has been an urgency to propound attack mitigation measures for smart grid. Honeypots,as an efficient deception defense method,can effectively collect attacks in smart grids. To address the issues of insufficient interaction depth,deficiency of physical industrial process simulation,and poor scalability in existing smart grid honeypots,this paper designs and implements a reinforcement learning-based smart grid honeypot framework—SGPot. It can simulate control side of a smart substation based on the system invariants in real devices of the power industry. Through the simulation of the power business process,the SGPot can enhance the deception of the honeypot and induce attackers to interact deeply with the honeypot. In order to evaluate the performance of the honeypot framework,this paper builds a small smart substation experimental validation environment. Meanwhile,SGPot,the existing GridPot and SHaPe honeypots are simultaneously deployed in the public network environment,and 30 days of interaction data are collected. According to the experimental results of this paper,the request data collected by SGPot is 20% more than GridPot and 75% more than SHaPe. SGPot can induce attackers to interact with the honeypot in greater depth than GridPot and SHaPe,and it obtains more sessions with interaction lengths greater than 6.

Keywords　Smart grid,Reinforcement learning,Intelligent interaction,Active defense,Honeypot

到稿日期:2022-11-22　返修日期:2023-02-20
基金项目:国家重点研发计划(2020YFB2010900);中原科技创新领军人才(224200510002)
This work was supported by the National Key R&D Program of China(2020YFB2010900) and Program for Innovation Leading Scientists and Technicians of Zhongyuan(224200510002).
通信作者:魏强(prof_weiqiang@163.com)

1 引言

随着信息系统和物理系统的不断融合发展，原本隔离的智能变电站电力设备开始直接或间接地接入互联网。由于智能电网安全措施薄弱且具有重要的军事、经济价值，其迅速成为了网络攻击的重要目标，面临着越来越大的网络攻击威胁。近年来，智能电网的攻击事件频发，造成了破坏性的影响。例如乌克兰电网攻击[1]，攻击者通过 BlackEnergy 建立据点，配置 build_id 的值来甄别受感染的目标，以“跳板机”为据点进行横向渗透，之后攻陷监控/装置区的关键主机[2]。模块 OPC.exe 发送一个 0×01 状态，这对于目标系统而言等同于“主变量超出限制”，会误导操作员认为是保护继电器状态，使操作员丧失可见性，从而操纵和控制断路器的开关。同样的事件还有印度核电站内网感染恶意软件[3]和委内瑞拉电力系统遭遇多次网络攻击导致大规模停电事故[4-5]等，都是先对电力电网进行侦察，找到突破点，与控制器建立连接，寻找关键 I/O 地址，从而利用一定技术手段骗过系统管理员，并实现业务流程的控制，以此达到攻击的目的。

蜜罐作为现有防御手段（如入侵防御系统和移动目标防御）的重要补充，是一种有效的欺骗防御手段，在智能电网的安全防护中发挥着至关重要的作用。通过研究智能电网蜜罐，能够对电力信息物理融合系统进行深度分析，探查可能存在的风险点位，为未来的电力设备等的设计提供指导。同时，研究智能电网安全防御技术，对提升智能电网的防御能力，以及保护关键工业基础设施安全稳定运行具有重要意义。目前，研究人员已经提出一系列的智能电网蜜罐，可以分为基于协议模拟的蜜罐和对设备/系统仿真的蜜罐。基于协议模拟的蜜罐通常会模拟设备的通用服务（如 HTTP(S)，SNMP，FTP 和 SSH 协议[6-7]）、智能电网数据传输协议（如 IEC 61850 GOOSE/MMS，IEC 60870-5-4 和 DNP3[8-11]）、用于编程控制的私有协议（如 S7comm 协议[8,10]）。此外，还有部分基于虚拟仿真软件构建的蜜罐，如基于 GridLAB-D 模拟电网的 GridPot[8] 和 Mashima 等[11] 基于 SoftGrid[12] 的研究。还有基于真实设备构建的蜜罐，如 DefRec[13]。同时，CryPLH[6] 和 Mashima 等[10] 的研究实现了对西门子 S7 系列可编程逻辑控制器（Programmable Logic Controller，PLC）的模拟，SHaPe[7] 实现了对符合 IEC 61850[14] 标准的智能电子设备（Intelligent Electronic Device，IED）的模拟。只进行通用服务和数据传输协议模拟的蜜罐缺少对设备和私有协议的模拟，虽然其能够通过攻击者的初步侦察，但是容易在深度交互中被识别，不能有效捕获后续攻击；而现有的模拟私有协议，同时结合了虚拟仿真软件和对设备仿真的蜜罐，针对电力物理工业过程的模拟不完善，定制化太强，扩展性差。

智能电网领域设计和部署蜜罐的难点在于需要对电力业务和工业过程进行模拟和仿真。具体表现为：1）仿真系统不全面，限制了蜜罐系统的真实性。目前的电力蜜罐大多是利用已有开源工控蜜罐和较易实现的广泛使用的 PLC 等的模拟，这部分体现出来的开源蜜罐特征和 PLC 指纹虽然能够有效模拟，但也存在会被识别的风险，且已有的方法在应用于智能电网中时对于电力设备执行控制指令等不能有效响应，这大大限制了蜜罐在电网中的应用。通过高仿真，可以提升蜜罐系统深度交互能力，提高欺骗性。2）构建高交互蜜罐需要的成本高，难度大。大部分高交互蜜罐的构建会结合或多或少的真实设备，而电力设备一般比较昂贵，导致成本高昂。因此构建基于真实设备与虚拟技术，能够不影响原有电力系统运行的高交互蜜罐，其实现和部署难度大，可扩展性差。

针对目前智能电网领域蜜罐技术的不足和挑战，本文提出了一种基于强化学习的智能电网蜜罐框架——SGPot。本框架具有以下特性：1）真实性和欺骗性更高。通过在真实智能变电站测试环境相同类型和配置的主机上绑定虚拟服务，SGPot 能够提供全面一致的真实视图，从而增强真实性，实现更高的欺骗性。2）深度交互能力进一步提升。该框架通过利用网络中公共可用的电力设备，学习其行为特征，收集设备对于蜜罐捕获请求的响应，对于当前蜜罐交互不深的问题，建模攻防行为，继而将其交互及响应决策建模为马尔可夫决策过程，并利用强化学习方法近似求解最佳响应。

为了验证 SGPot 的有效性，本文进行了一系列功能验证和性能测试实验。功能验证实验用于验证 SGPot 在欺骗性、信息要素收集完整性等方面的优越性。性能测试方面主要用于验证深度交互和业务仿真的能力，具体地，本文将 SGPot 和现有的开源蜜罐 GridPot 和 SHaPe 部署在云服务器上，进行了为期 30 天的请求数据采集，SGPot 共收集到 1 678 个请求，对采集数据进行分析的结果表明，SGPot 能够通过智能交互欺骗攻击者深入系统，提供与真实电力系统相当的交互性，在深度交互上优于现有方案。此外，本文还对比了真实场景和 SGPot 在执行相同电力操作时的响应时间，可以看出其均在正常范围内，且与真实环境时间几乎一致。

本文的主要贡献包括以下 3 个方面：

1）设计和实现了电力物理工业过程的模拟，能够有效通过攻击者对智能电网的初步侦察。

2）提出了一种基于强化学习的智能电网交互框架，设计了框架的多模块和信息流的传递，能够及时有效地提供攻击者预期的响应，从而进一步扩展与攻击者的会话，欺骗攻击者在系统中停留更长时间，分析其攻击意图，从而达到保护真实电网设备/系统的目的。

3）在搭建的小型智能变电站环境中进行了实验验证，证明了框架的有效性。与已有开源蜜罐相比，所提方案提高了与攻击者深度交互的能力。

2 相关工作

蜜罐是一种通过工具诱骗攻击者，令安全人员得以观察攻击者行为的主动防御技术，其更多地关注攻击者本身。通过欺骗诱捕打乱攻击节奏，提升攻击复杂度，蜜罐能够给防御方增加更多响应时间，并有可能对攻击者进行分析溯源，从而预防攻击。本节对智能电网相关蜜罐和基于机器学习的自适应和智能蜜罐研究的现状进行了分析。

2.1 智能电网蜜罐/蜜网

对于智能电网的蜜罐研究，目前仍然处于起步和探索阶段。已有的研究从最初的只针对协议模拟到后来提高仿真度和业务过程的模拟和仿真，逐步提升了防御效果。主要的一些智能电网蜜罐/蜜网研究如表 1 所列。

Buza 等[6]为了检测针对工业控制系统(Industrial Control Systems，ICS)的针对性攻击，设计并实现了 CryPLH，它模拟了西门子 Simatic 300 PLC，能够有效捕获该类 PLC 的攻击，但是可扩展性差。SHaPe[7]模拟了任何符合 IEC 61850 标准的 IED，是第一个为 SCADA 网络设计的 Dionaea 模块，其利用 Libiec61850 库处理制造报文规范(Manufacturing Message Specification，MMS)请求，能够进一步在协议层提高交互性，但是没有对工业过程进行模拟，在交互过程中不能及时提供正确的响应。Redwood 等[8]提出了一个符号化的信息物理蜜网框架 SCyPH，其利用系统物理的符号数据流模型，通过集成物理模拟和异常检测扩展了蜜网。他们用一个电网来演示 SCyPH 框架，即 GridPot，利用 GridLAB-D 模拟器进行变电站模拟和基于 IEC 61850 的通信，且利用 Conpot[15]来模拟 IED，并为设备之间的交互实施了 GOOSE/MMS 和 Modbus 协议。由于加入了电网的模拟，其仿真度进一步提高，但是在设备的模拟上较单一，且扩展性较差。

表 1 智能电网蜜罐/蜜网

Table 1 Smart grid honeypots/honeynets

Honeypot	Year	Level of Interaction	Simulation service/technology	Open sources
CryPLH[6]	2014	Low	HTTP(S)，SNMP，Step 7	×
SHaPe[7]	2015	Low	IEC 61850 MMS，HTTP，FTP，SMB	√
GridPot[8]	2015	Hybrid	IEC 61850 GOOSE/MMS，Modbus，HTTP；GridLAB-D；IED	√
Mashima 等[9]	2017	Low	IEC 60870-5-104，IEC 61850，SSH	×
Mashima 等[10]	2019	Low	TCP port listener on IEC 61850 MMS，IEC60870-5-104，EtherNet/IP，DNP3，Modbus/TCP，S7comm，Niagara Fox，BACnet	×
Mashima 等[11]	2020	Hybrid	IEC 60870-5-104，IEC 61850 MMS，Back end power flow simulator，SSH，VPN，SoftGrid	×
DefRec[13]	2020	Hybrid	PFV，SDN，(Real devices)	×

Mashima 等[9]设计并实现了一个智能电网蜜网系统，利用虚拟化技术模拟包括多个联网变电站在内的整个智能电网现场通信基础设施。Mashima 等[10]部署了低交互智能电网蜜罐系统，基于观察的攻击来调整真实系统，但是没有对高保真工控蜜罐进行分析。Mashima 等[11]设计了一个智能电网通信基础设施蜜罐，并从攻击者的角度对智能电网蜜罐系统进行了评估，进一步提高了其真实性。这 3 类蜜罐逐步提高了蜜网的模拟程度，真实性得到进一步提高，但是针对性强，环境构建难，成本高。Lin 等[13]为了干扰攻击者对智能电网等 ICS 的侦察，提出了物理功能虚拟化，将网络交互与真实设备“挂钩”，并使用真实设备构建轻量级虚拟节点，这些节点遵循了真实设备中网络堆栈、系统不变量和物理状态变化的实际实现。之后，在物理功能虚拟化基础上，他们提出了 DefRec，该防御机制能大大增加攻击者推断电网网络物理基础设施的难度。但是，其采用的是真实节点，成本高。

2.2 基于机器学习的典型蜜罐研究

随着人工智能、大数据的快速发展，机器学习受到了越来越多的关注，被各行各业用于实现高端和智能技术的开发。攻击者使用的工具也进一步趋于多样和智能化，这就需要以智能对抗智能的方式来设计和实现蜜罐，从而自适应和智能地与攻击者交互以捕获高价值的攻击和零日漏洞。表 2 列出了一些基于机器学习的典型蜜罐。

表 2 基于机器学习的典型蜜罐

Table 2 Typically honeypots based on machine learning

Honeypot	Year	Level of Interaction	Methods and Technology	Open sources
Wagener[16]	2011	self-adaption	Reinforcement learning，Game Theory	×
IoTCandyJar[17]	2017	Intelligent	Reinforcement learning	√
Pauna 等[18-20]	2014 2018 2019	self-adaption	Reinforcement Learning，Deep Q-Learning，NN，Inverse Reinforcement Learning	×
FirmPot[21]	2021	Intelligent	Attention mechanism，RNN，Word2Vec	√

Wagener[16]利用博弈论概念定义高交互蜜罐的配置和动作。其使用强化学习的一种变体，在面对攻击者时获得了最佳行为。其主要针对 SSH 服务器的相关工作，对于响应定义有限，不能直接应用于工业场景中。Luo 等[17]使用机器学习技术自动学习物联网设备的行为并构建“智能交互”蜜罐，利用互联网上公共可用的物联网设备来收集蜜罐捕获请求的潜在响应，能够获得不同类型物联网设备的行为，并利用多种启发式和机器学习机制来自定义扫描程序并改进回复逻辑，从而更大程度地扩展会话。由于工业互联网和物联网在设备等方面存在较大差别，该方法在工业场景中的信息收集和潜在响应的收集不能有效达成。

Pauna 等开发了一种自适应中等交互的蜜罐系统

RASSH[18]来模拟 SSH 服务器，能够通过强化学习算法与攻击者进行交互。之后，他们开发了 QRASSH[19]，使用深度 Q 学习算法来决定如何与外部攻击者交互，集成了利用神经网络的强化学习算法的现有实现。而后，他们进一步对其进行改进，实现了一个 SSH/Telnet 蜜罐系统[20]，并给出了定义要使用的最佳奖励函数的结果。以上 3 个蜜罐采用强化学习方法，但是聚焦于 SSH 等服务的研究，对工控协议及场景不能适用。

Yamamoto 等[21]设计并实现了一个使用固件自动生成物联网蜜罐的框架 FirmPot。使用一个针对蜜罐生成优化的固件模拟器，并通过机器学习来学习嵌入式应用程序的行为。FirmPot 通过模拟在 docker 容器上启动的固件来收集 Web 交互，并通过机器学习来学习请求和响应的对应关系。此外，生成的蜜罐会从模拟的响应中返回对接收到的请求的最佳响应。该方法聚焦于物联网，其固件已经有很多工作进行研究，固件获取较易，但是在电力行业中，设备及系统等多样异构，固件几乎无法直接获取，因此也不能应用于电力行业的防御。

3 问题挑战

因为电力环境具有独特的功能，使得包括蜜罐在内的安全工具很难在这些领域高效部署。SCADA 等电力设备/系统需要保证连续工作，并高度避免中断和停机，且其中的设备通常需要严格执行响应时间的实时性。由于这些原因，在智能电网环境中插入蜜罐非常困难，可能会影响电力通信并有被入侵的风险。只有中高交互蜜罐才有可能看到其他更高级的攻击和工业协议并处理特定的攻击，以及才能识别可能针对电力设备的复杂攻击，并了解对工业流程和关键基础设施的可能影响。而蜜罐允许的交互级别越高，它可能被攻击者破坏和使用以损害网络中的其他系统，甚至对其他网络发起攻击的风险就越大。

在部署高交互蜜罐时必须非常小心，尤其是在智能电网监视和控制关键工厂流程的环境中部署蜜罐时需要确保它们不会被攻击者破坏，并确保它们不会干扰现有电力设备的通信和控制过程。

因此，目前的挑战在于从网络流量、连接设备等中捕获和过滤这些攻击痕迹，并开发出能够有效利用这些数据来破坏正在进行的攻击并防止未来企图执行的攻击的防御机制。从攻击链的角度分析蜜罐的设计，根据其设计要素，目前在全要素信息收集、业务流程的模拟和智能交互方面面临挑战。

3.1 全要素信息收集的挑战

电力工控系统中大量使用智能设备、嵌入式操作系统和各种专用协议，具有集成度高、行业性强、内核不对外开放、数据交互接口无法进行技术管控等特点。在进行信息收集的过程中，需要筛除掉网络中蜜罐等不真实设备的流量，同时要覆盖尽可能多的设备。智能电网相对封闭，攻击者或是有组织的、长期的 APT 攻击，或是针对特定行业具有行业知识，对系统等非常熟悉，个人攻击者很少，因此收集数据信息具有挑战。

3.2 业务流程仿真的挑战

智能电网中构建真实可行的蜜罐需要以近乎真实的业务流程来支撑，而智能电网区别于其他 ICS，其逻辑、架构完全不同。在水处理系统中，可以基于 minicps[22]和 mininet[23]进行水塔的仿真，借助仿真的 s7 系列 PLC 来仿真控制。而智能电网中，需要结合电力运营、潮流分析、总线系统等进行业务流程的仿真。其次，需要相关架构知识作为背景。因此，对智能电网业务流程进行有效的仿真具有挑战性。

3.3 智能交互的挑战

传统的蜜罐设计主要构建低中高交互类型，而低交互不能捕获高级攻击，高交互结合设备，构建成本高。现有研究几乎都是低交互的，具有固定的回复逻辑和有限的交互级别，无法捕捉到有价值的攻击。因此需要通过学习设备或者系统行为，基于强化学习的方法来进行智能交互。而在智能电网中学习设备行为特征具有挑战性。因为在实际操作中，给设备发读/写指令后需读/写的位置不同，返回的响应也不同，同时需要真实响应。例如，考虑测控装置的变化，在监控端的操作需要实时在测控装置有相应的反应。

4 SGPot 设计与实现

为了应对以上挑战，本文提出了一个针对智能电网的蜜罐框架(见图 1)，用于为智能电网电力系统构建智能交互蜜罐。

智能交互的目标是从关于智能电网实体设备的“零知识”中学习与攻击者交互的“正确”行为，初始时 Agent 不掌握攻防知识和行为分析数据，只进行随机的响应选择来传回蜜罐实例以应对各类攻击者，通过对交互的深入判定和是否有攻击行为的检测来认定响应是否有效，从而以最初的随机化模式逐渐学习到设备行为特征，在后续交互中生成攻击者期望的和基于设备的正确响应。在整个过程中，Agent 均基于其针对电力设备行为学习到的特征来进行响应的生成。对攻击者的正确响应应该能够扩展与攻击者的会话，以欺骗他们通过最初的侦察并执行后续攻击操作。为了实现这个目标，需要通过构建的系统自动收集有效的响应作为候选。通过不断地与攻击者交互，决策模块能够为蜜罐实例提供针对每个请求的正确行为并不断优化。

SGPot 框架主要由 5 个部分组成，运行过程中定时相互共享数据。知识库包括一个记忆库和一个攻防知识库，用于存储获得的所有相关的智能电网电力设备信息等。智能决策模块利用强化学习方法，根据攻击者交互的请求响应来训练模型，经过多次学习，可以优化决策一个响应供蜜罐来回复攻击者。信息收集模块利用捕获的请求作为种子，并扫描互联网以查找可以响应这些请求的任何电力设备，同时将请求发送到测试环境进行响应的收集。收集到的响应存储在原始响应表中，并存放在记忆库中。

本文工作的目标是建模、实施和评估适应攻击者的智能

蜜罐框架，以获悉更多关于攻击者的知识。本文介绍了对攻击者行为的建模以及适应的可能性，通过构建自适应蜜罐，自动利用这些特征来应对互联网上的各种攻击。强化学习让蜜罐更接近通过增益函数参数化的最佳行为。

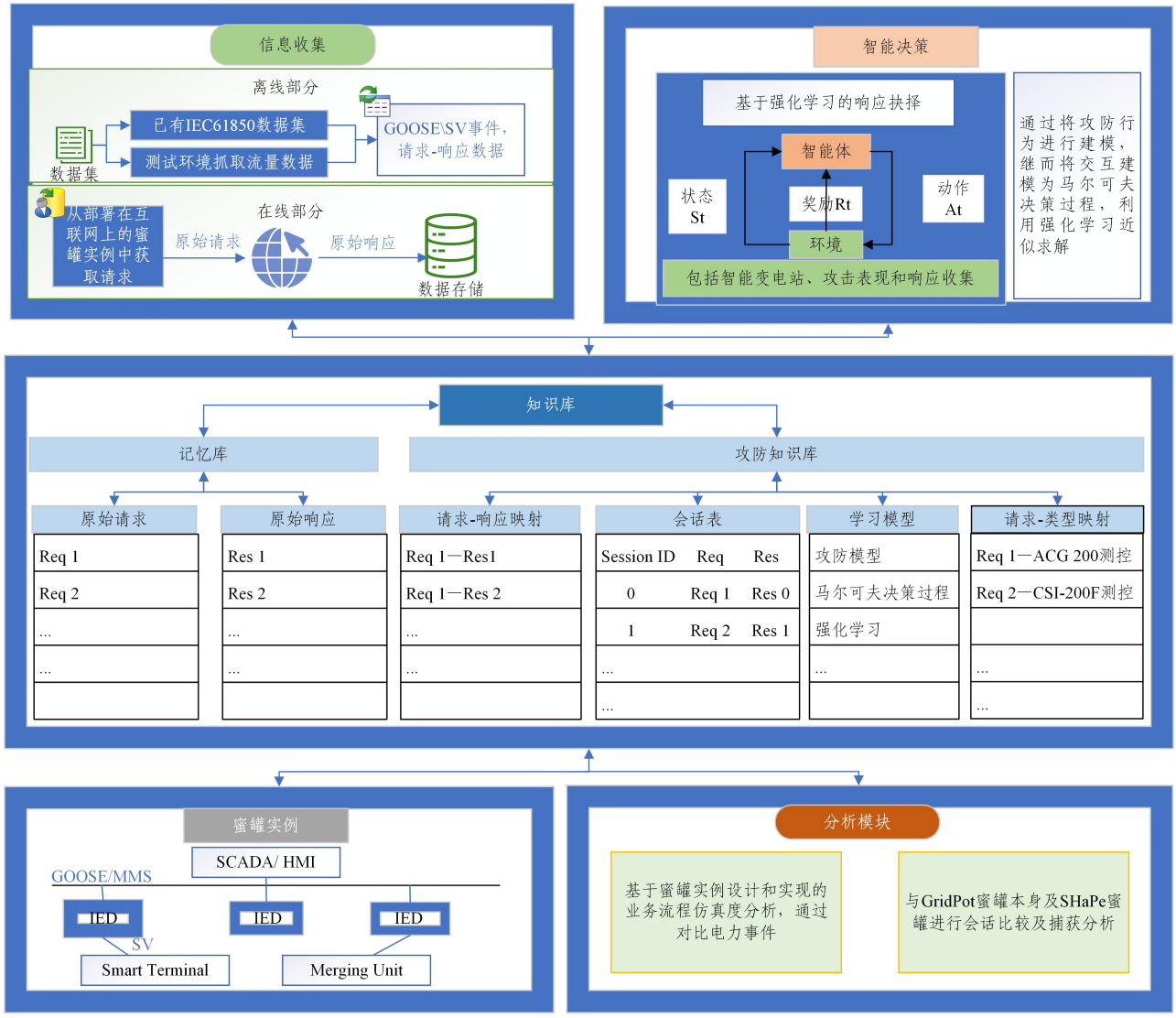

图 1 SGPot 整体架构

Fig. 1 SGPot overall architecture

4.1 信息收集

为了应对全要素信息收集面临的挑战，本文设计并实现了多维全要素信息收集模块。由于电力行业具有特殊性和保密性，环境相对封闭，设备根据特定行业多样异构，导致信息获取较难。但是调查分析显示，目前联网电力系统中暴露在公共互联网的电力行业网络资产趋于上升态势，有数据显示 2 个季度收集的资产就多达 1 000 多个，主要是设备资产，还有部分传统电力 web 资产和新能源智能电站及 web 资产。重点是 IEC-104，Modbus 等电力协议相关的端口被广泛探测。因此，对联网公开暴露的电力系统的信息收集可以达到目标。

采用离线加在线的方式全要素获取相关信息。首先，在离线阶段，基于已有数据集和测试环境进行流量分析，提取请求和响应要素，对设备及参与会话个体进行表征。同时，采用在线获取的方式，在 Shodan 和 FOFA 平台进行了测试，发现在获取远程终端单元（Remote Terminal Unit，RTU）等设备的信息时，几乎全部都被标记为蜜罐，这就意味着从真正真实的设备获取有效的响应是困难的，而 IED 可以获取正常的一些响应。在 Shodan 等平台进行信息收集，收集了 50 多个 IP 地址。将蜜罐实例部署在互联网上后，基于其被探测扫描的请求数据，将其转发到测试环境和在线可用的真实设备进行响应的收集。

4.2 蜜罐实例

蜜罐实例针对智能电网业务流程的仿真，解决了业务流程仿真的挑战。

如图 2 所示，蜜罐的部署主要基于改进的 GridPot，将 GridPot 基于已搭建实体设备环境改进其 SCADA 系统界面，相关界面和信息等采用真实完整的显示，将其部署于验证环境的监控站 SCADA 主机中，相当于同时有两个 SCADA 系统，一个是原来的系统，一个是蜜罐实例界面，蜜罐实例提供全面一致的视图，主要完成获取真实攻击者的请求和请求的真实响应。

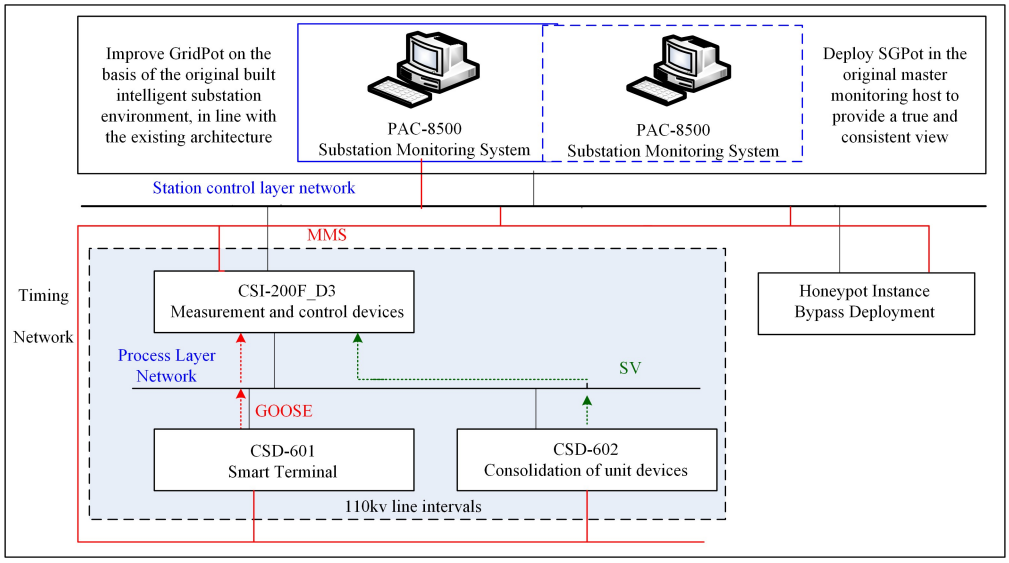

图 2 蜜罐实例框架图

Fig. 2 Framework diagram of honeypot instance

如果攻击者没有识别到 SCADA HMI 是假的，那么他就会在虚假的系统中继续寻找可突破的点；如果攻击者识别到是虚假的 SCADA HMI，则会终止交互。同时，如果攻击者的攻击载荷或者恶意软件足够智能化，蜜罐在与攻击者交互的过程中没有及时捕获，误判了攻击者的意图，那么攻击者很大概率会将其识别为蜜罐，从而很可能会终止操作。

4.3 智能决策

通过将攻防行为建模，由于其交互过程中系统操作者主要完成最佳响应的选择，因此将最佳响应的抉择问题进一步建模为马尔可夫决策过程。之后利用强化学习的方法求解近似最佳响应，从而针对性地为智能交互的挑战提供了解决方案。

4.3.1 攻防模型的建立

本节从智能变电站可能受到攻击的角度分析攻击者行为及攻击链的预测。如图 3 所示，基于单一小型智能变电站，攻击者可能发起两类攻击。首先，攻击者需要通过互联网探测扫描蜜罐系统，通过渗透的方式对 SCADA 主机发起攻击，其主要从控制中心端和过程层网络实施攻击。

攻击 A：针对 SCADA 系统控制中心的攻击。这类攻击可能破坏电网系统的完整性。一方面，攻击者足够强大，其通过一系列措施获得了 SCADA 系统的访问权，然后就可以向变电站发出任意的、看似正常和合法的控制命令，可能造成一系列错误决策，导致片区停电等事故。另一方面，攻击者只获得了 SCADA 系统中正常工作站的访问权，之后，攻击者可能通过中间人攻击发送重放或者伪造的命令，还可以拦截正常来自变电站的消息。

攻击 B：一些变电站有意或无意连接到了互联网公网，这为攻击者提供了入口点。同时，即使是专用的基础设施也可能不是完全安全的，例如，对于在边界连接的智能设备，网络设备的中继端口等可能被利用。无线或蜂窝通信的方式使得攻击者可以通过损害中间基站来实施中间人攻击。此外，攻击者还可以建立流氓基站，以误导通信，并冒充控制中心发送错误指令。

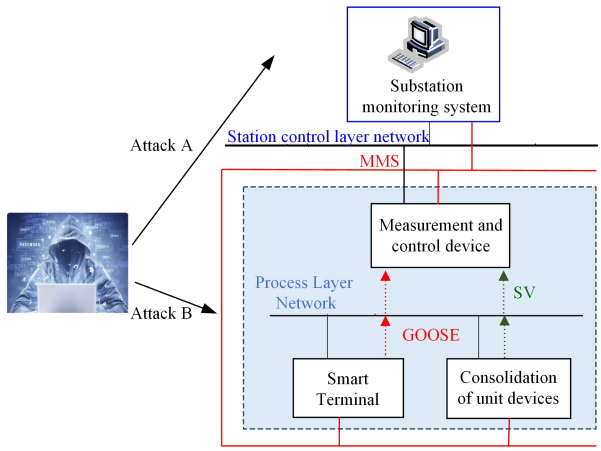

图 3 测试环境智能变电站典型攻击模型

Fig. 3 Test environment intelligent substation typical attack model

为了进一步对攻击者和系统操作员行为进行建模，本文从攻防博弈的角度入手。攻击者和防御者之间的较量和交互本身就是一种博弈，由于资源的不对称，系统操作者不知道什么时候什么节点会有什么样的攻击发生，且攻击态势呈现组织模式，因此单个攻击者行动很少。

攻击者的目的是最大化攻击效果，尽可能多地窃取威胁情报和电网的网络拓扑，执行攻击以达到大停电等事故。而防御者要最小化影响，在已有防御措施的基础上，采取一定的行动以反制或者欺骗攻击者，从而防御真实的设备或系统。

在交互过程中，一方面要捕获攻击，分析攻击者意图；另一方面要记录攻击者的活动路径，探查攻击者针对的目标。图 4 给出了攻击链模型。由于最初无法判定攻击者可能最终执行的路径，因此用实线代表最终攻击者的路径，虚线代表攻击者在执行最终路径时做过的其他攻击

尝试。其中红色代表攻击者在确定攻击目标之后利用远程桌面登录等方式对 SCADA 系统所在主机进行的初步侦察和探测扫描活动，从而尽可能详尽地描绘智能电网网络拓扑以及可能的设备型号等，攻击链中会记录攻击者的探测扫描日志、攻击者的侦察范围和侦察效果等。浅绿色代表攻击者在初步的侦察中，没有识别到目标对象是虚假的，并且在侦察中发现了足够的网络拓扑信息并能够提取部分攻击特征，对正在运行中的终端设备进行漏洞发现和基于在初步探测中在 SCADA 系统中获取的 IP 等信息，利用白名单介入方式使得自己可以连接和修改部分信息，例如发现设备 SSH 和 TFTP 等服务端口的开启情况、初始密码等。之后，攻击者利用手段控制 SCADA 系统或者从终端设备中发现可利用的点来执行下一步的攻击。攻击链在这一步的构建中，需要详细记录攻击者从上一节点如何到达该节点，是否植入攻击载荷，或者采取了哪些攻击等。紫色则是代表攻击者在进一步深入交互过程中可能执行的路径和采取的方式等。

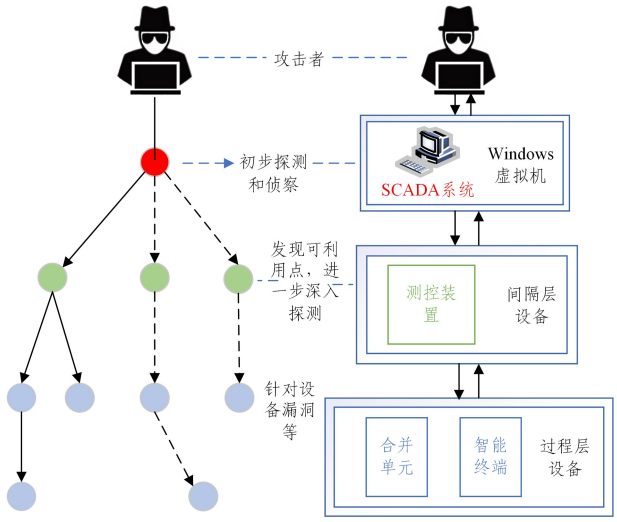

图 4 攻击链模型(电子版为彩图)

Fig. 4 Attack chain model

电网正常运行的情况下，攻击者对电网的任何操作都要在其正常范围内，才能避免不被已有的检测机制检测到，因此限制攻击者操作的问题有：对电力电网各类数据的正常范围的考量、电力网络拓扑结构的获取以及对各类设备响应时间的判断等。

智能电网电力传输中主要涉及发-输-变-配-用 5 个环节，如图 5 所示。其中，发电端产生电能，经输电线路将电能传输到远方。为了避免电能的浪费，需要将电能电压等级进行调节，之后在配电站将电能分配给用户。整个传输线路的电能传输过程中各个环节都有可能遭受网络攻击。

由于攻击者在实施攻击的过程中，其行为在整个攻击过程中是连续的，但是具体操作在时间上可能是不连续的，例如，在乌克兰电网攻击事件中，攻击者需要匹配攻击特征 ID。因此，攻击者的行为模型可以表述为离散时间动态系统。

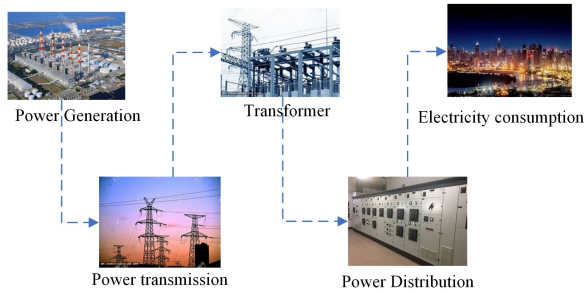

图 5 智能电网电力传输示意图

Fig. 5 Schematic diagram of smart grid power transmission

对于攻击者在某一时刻 t 实施的攻击 x，可以表示为：

$$x(t)=x, x\in X \tag{1}$$

其下一步要执行的操作与该阶段实施的攻击效果相关，如果初步会话达到了预期的效果，例如侦察过程中发现了感兴趣的点，攻击者就会针对该点进一步探测是否存在漏洞等情况。如果在初步的操作中没有得到预期的结果，则可能会终止此次行动。具体表达式如下：

$$x'(t)=\begin{cases} x_i, i=0,1,2,\cdots,n, x\in X, \\ \qquad \text{Get the expected response} \\ 0, \qquad \text{Didn't achieve the excepted goal} \end{cases} \tag{2}$$

由于攻击者采取的任何行动不能对系统产生较大影响，以免被系统检测到异常，因此攻击者采取的行动需要满足在系统变化的正常范围内，这个约束可以表示为：

$$-\omega x(t)\leqslant \Delta x(t)\leqslant \omega x(t),\ \forall x(t)\in X \tag{3}$$

其中，ω 是一个参数，用于表征前后量的变化范围。攻击者的目标是最大化攻击效果，在窃取到足够多的威胁情报的同时最大化破坏力，使目标网络和设备瘫痪，造成级联故障或者大范围停电事故。

与此同时，系统操作者需要及时捕获到攻击者的行为，从而进行预判，来做出反应。根据典型的安全约束经济调度模型来看，系统操作者受到以下条件的约束：

系统负荷平衡约束：

$$\sum_{i=1}^{I} P_i(t)=P_d(t),\ t=1,2,\cdots,T \tag{4}$$

其中，$P_d(t)$ 为系统 t 时间段的总负荷。

电网安全约束为：

$$|P_{ij}(t)|\leqslant \bar{p}_{ij} \tag{5}$$

其中，p_{ij} 和 $\bar{p}_{ij}$ 分别为支路 $i-j$ 的潮流功率及上限。

除此之外，系统操作者还会受到电网运行约束和现有可用防御技术的限制以及可操作的空间。同时，系统操作者回复响应的时间要在电力设备/系统正常时间范围内，可以表述为：

$$t-\mu\leqslant t\leqslant t+\mu \tag{6}$$

其中，μ 是一个时间度量。

攻击者在真正实施攻击前会做好准备，因为每一次不成功的尝试都可能导致其身份暴露，揭示其攻击方法等。如果攻击者是像执行 APT 一样或者想要达到级联攻击效果，则其会在前期侦察过程中花费较多的时间来布局攻击，以获取到

足够多的威胁情报数据，使得攻击目标透明化；而执行协同攻击的攻击者则需要针对不同的节点，采用相同或者不同的技术手段来达到攻击的效果。

4.3.2 马尔可夫决策过程模型

针对智能电网中的针对性攻击的有效防御策略之一是欺骗防御技术。欺骗防御技术通过设计诱骗系统，诱导攻击者将目标转向虚假系统，从而保护真实设备/系统。而系统操作者在与攻击者交互的过程中如何选择一个合适的攻击者预期的响应进行回复，诱导攻击者进一步对所在的系统进行探索成为了解决该问题的关键。假设攻击者是继续会话还是执行攻击仅由前一个请求的响应来决定，基于之前的研究和交互的随机性，这个假设是合理的。马尔可夫链是满足马尔可夫性质的随机过程，将来的状态只取决于现在，与过去无关。因此将最佳响应的选择决策问题建模为马尔可夫决策过程(Markov Decision Process，MDP)，以求解最佳响应，欺骗攻击者长时间驻留在系统中，从而有效捕获攻击者行为，分析其路径和攻击目标。

MDP 是马尔可夫链的一种扩展，它提供了一个用于对决策情景建模的数学框架。系统操作者可以被视为一个 Agent，旨在在不确定的环境中做出序列决策，通过不断的响应回复来加深与攻击者的交互，最大化累计奖励，从而捕获高级攻击。Agent 需要与环境连续交互，在每个阶段的交互中回复响应，并记录攻击者的活动，以分析其意图。更具体地说，用以下 5 个关键要素表示 Agent 的决策过程。

状态 S：由于攻击的发现和被识别的必然性，攻击者执行攻击在系统中所能存在的状态是有限的。

动作 A：在每个阶段，Agent 选择一个响应回复给攻击者，并记录攻击者该阶段的状态和使用的手段等。由于交互过程会在某个阶段终止，同时基于响应集的有限性，动作空间 A 也是有限的。

转移概率 P：系统的动态由转移概率决定，它表示 Agent 回复一个响应执行一个动作后，从当前状态到下一个状态的状态转移。

奖励 r：Agent 执行某一动作 a，从一个状态 s 转移到另一个状态 s'所获得的奖励。确定性奖励 $r(s,a)$与状态-动作对 (s,a)相关。

折扣因数 λ：折扣因数控制着即时奖励和未来奖励的重要性。折扣因数的值位于 0～1，折扣因数为 0 意味着即时奖励更重要，而折扣因数为 1 则表明未来奖励起到了关键作用。

执行动作 a 从状态 s 到 s'的转移概率表示为：

$$P_r(s'|a,s)=pr(s_{t+1}=s'|s_t=s,a_t=a) \tag{7}$$

执行动作 a 从状态 s 到 s'所得到的奖励表示为：

$$R(s'|a,s)=\mathrm{E}(R_{t+1}|s_t=s,s_{t+1}=s',a_t=a) \tag{8}$$

Agent 试图使得从环境中获得的总奖励(累计奖励)最大化(称为回报)。因此，Agent 所获得的回报如下：

$$R_t=r_0+r_1+\cdots+r_t \tag{9}$$

其中，r_{t+1}是 Agent 在执行动作 a 从一个状态转移到另一个状态过程中在时间步 t_0 所获得的奖励。在具体实现中，即时奖励反映了当选择响应 a_t 来回复请求 Req_t 并移动到下一个状态 s_{t+1}时在交互中取得的进展。由于进度可以是正的也可以是负的，因此奖励也可以是正的或负的。在执行交互过程中，如果选择的响应正好是攻击者预期的响应，并且攻击者继续执行以部署攻击，则其分配的奖励是正的且较大，相反会导致负奖励。将奖励分配为等于最终会话长度的值，因此可以认为交互的请求越长，包含恶意负载、攻击载荷的概率就越大。

综合以上定义可知，MDP 中，决策过程由 Agent 可以访问的一组状态组成。在每个状态下，Agent 都可以在一组动作中选择一个动作。奖励函数定义了在每个动作之后分配的奖励。Agent 搜索由状态和动作之间的关系定义的策略，目的是通过最大化收益来找到策略。通常，MDP 中的转移函数以及奖励函数是未知的。此外，贝尔曼方程的计算复杂度非常高。这两个问题使我们转向强化学习，从而近似最优策略。

4.3.3 基于强化学习的最佳响应的判决

强化学习模型中，Agent 与环境交互，采用 Agent 感知环境的方式，基于输入选择要在环境中执行的动作，通过与动态环境的试错来学习从状态到动作的映射关系。

其中，策略函数可表示为 $\pi(s):S->A$，这表示从状态到行为的映射。策略函数表示在每个状态执行什么动作，其目标是找到在每个状态指定正确行为的最优策略，从而使奖励最大化。

状态值函数确定一个 Agent 在策略 π 下处于某一特定状态的最佳程度，通常记为 $V(s)$，表示执行策略后状态的值。可以定义为：

$$V^{\pi}(s)=\mathrm{E}_{\pi}[R_t|s_t=s] \tag{10}$$

状态-行为值函数，即 Q 函数，用于表明智能体遵循策略 π 在某一状态所执行的特定行为的最佳程度。Q 函数的定义如下：

$$Q^{\pi}(s,a)=\mathrm{E}_{\pi}[R_t|s_t=s,a_t=a] \tag{11}$$

式(10)确定了根据策略 π 从状态 s 开始采取行动 a 所获得的期望回报。

基于此，最佳响应的选择问题就转化为寻找最优策略和值函数，通过取 Q 函数最大值可以更简单地计算最优值函数：

$$V^{*}(s)=\max_{a} Q^{*}(s,a) \tag{12}$$

因此，Q 函数的贝尔曼方程可以表示为：

$$Q^{\pi}(s,a)=\sum_{s'}P_r(s'|a,s)[R(s'|a,s)+\gamma\sum_{a'}Q^{\pi}(s',a')] \tag{13}$$

将式(13)带入式(12)得到贝尔曼最优方程：

$$V^{*}(s)=\max_{a}\sum_{s'}P_r(s'|a,s)[R(s'|a,s)+\gamma\sum_{a'}Q^{\pi}(s',a')] \tag{14}$$

QL 算法的目标是根据贝尔曼方程更新 Q 值函数，通过求解贝尔曼最优方程，可得到最优策略，从而为 Agent 找到最佳响应。

采用值迭代的方式找到最优策略：首先从一个随机 Q 函数开始，以迭代方式寻找一个新的改进 Q 函数，直到找到

最优值函数。一旦得到最优值函数，就可以从中得到最优策略。

首先，初始化随机值函数，即每个状态的随机值；然后，计算所有状态-行为对的 Q 函数 $Q(s,a)$；最后，用 $Q(s,a)$ 的最大值更新值函数。重复以上步骤，直到值函数的变化非常小。

根据下列方程更新 Q 值：

$$Q(s,a)=Q(s,a)+\alpha(r+\gamma \max Q(s',a')-Q(s,a)) \quad (15)$$

其中，$Q(s,a)$ 为状态-动作对 (s,a) 的 Q 值函数。α 表示控制学习响应的学习率，当 α 为 0 时，Agent 只会坚持最初的估计，不会从自己的行为中学到任何东西，因此 Q 值不会更新。当 α 为 1 时，Agent 只考虑最近的信息，这会导致快速学习，但是会丢失之前学到的知识。式(15)是一个自我更新的过程，可以迭代求解。

根据上面的描述，如图 6 所示，基于 QL 的响应判决可以指定为：Agent 从候选响应集中选择一个响应 a，然后根据当前状态计算是否诱导攻击者进一步深入交互来衡量响应的有效性(即奖励)$r(s,a)$。一旦选择了一个响应，系统状态将被更新，同时攻击者攻击行为被添加到攻击链。这个过程一直往下执行，直到攻击者出于某种原因终止攻击或者在交互过程中已经完全获悉攻击者针对的目标、采用的战术和工具等。

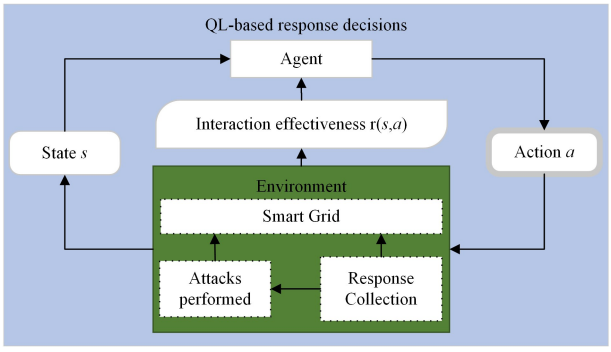

图 6 基于 QL 的响应判决

Fig. 6 QL-based response judgments

响应选择的过程如下：在选择响应进行回复时，如果 Agent 选择响应只是为了优化下一个即时奖励，那么响应判决过程可能会陷入局部最优。为了避免发生这种情况，Agent 对未知的请求的响应进行探索，通过执行 ε 贪婪策略来实现当前状态下以较大的 Q 值获得最大的奖励。其中 ε 表示 Agent 选择探索新的非最佳响应的小概率，状态 s 下的 Agent 所采取行动 a 的概率应满足：

$$P_r(a)=\begin{cases}1-\varepsilon, & a=\arg\max\limits_{a\in A} Q(s,a)\\ \varepsilon, & a\neq\arg\max\limits_{a\in A} Q(s,a)\end{cases} \quad (16)$$

首先生成一个随机数 ε，ε 足够小。Agent 以 1−ε 根据下一状态的 Q 值选择最佳动作，即 $\arg\max\limits_{a\in A} Q(s,a)$；以概率 ε 随机选择动作。算法 1 描述了提出的基于 Q-Learning 的最佳响应的决策过程。

算法 1 基于 Q-Learning 的最佳响应的决策过程

输入：Req

输出：a(即 Res)

1. Initialize π(s) arbitrarily
2. Initialize Q(s,a) arbitrarily for each s∈S,a∈A
3. Repeat(for each episode):
4. Initialize s
5. Repeat(for each step of episode):
6. Choose a from s using(16)
7. Take action a,observe r,s′
8. Update π using(14)
9. Update Q(s,a) using(15)
10. s←s′
11. until s is terminal
12. End

4.4 知识库

知识库分为记忆库和攻防知识库。记忆库存放原始请求和响应，以及大量的网络数据包，后经处理存放到攻防知识库中。攻防知识库中存放请求-响应映射、会话表、学习模型和请求-类型映射。请求-响应映射是记忆库中匹配的一系列关系；会话表是在智能交互过程中基于会话 ID、请求 ID 和响应 ID 等要素组成，便于通过 ID 检索到需要的会话模式；学习模型是在智能交互过程中实现的 MDP 推理及攻防行为的建模等。请求-类型映射是基于交互过程中生成的 MDP 状态图，从中抽取出请求和设备类型之间的关系。知识库中所有要素在交互过程中，与蜜罐实例、分析模块、信息收集模块实时同步信息。

知识库的建立能够帮助整合所有相关的信息，其中在 MDP 中生成的图信息可以在后期指导信息收集，同时为学习和攻防的推理以及攻击的构造等提供数据支撑。

4.5 分析模块

分析模块负责分析所有监控到的日志，例如它输出接收请求的统计信息以及与攻击者的会话长度。主要对捕获的攻击等行为进行分析，以挖掘攻击者真正想要攻击的目标或者下一个阶段要实施的攻击，为系统的保护提供数据支撑。

5 实验评估

5.1 实验设置

智能变电站测试床由仿真测试环境和实体设备组网测试环境组成，分别如图 7、图 8 所示。之后分别在两个测试环境中进行了相关实验验证。由监控站 SCADA 主机和测控装置、智能合并单元等构成了小型组网变电站测试环境，基于此环境能够开展协议分析，以及相关功能性验证。同时可以结合真实设备构建蜜罐实例。在实施过程中，将蜜罐实例部署在 SCADA 主机中，进行框架的功能性验证，同时部署在互联网中进行性能分析，我们统计了 30 天的数据。

本文实验主要从功能和性能两方面进行评估。功能性实验主要从信息要素收集的全面性以及构建诱饵系统框架的欺骗性两方面进行分析，以验证框架功能是否齐全。性能方面主要对会话的扩展和业务仿真的有效性进行评估，以验证在智能电网环境中是否达到预期的效果。

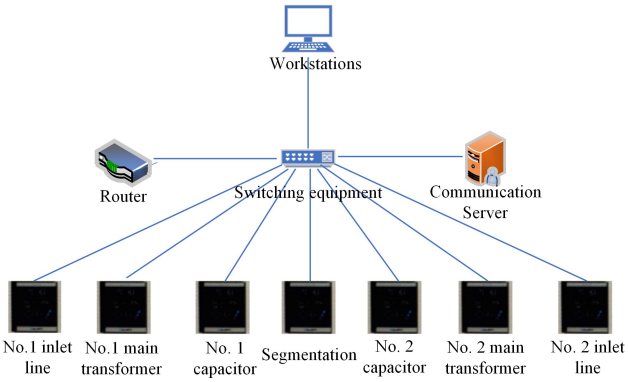

图 7 智能变电站仿真环境

Fig. 7 Smart grid simulation environment

图 8 实体设备组网测试环境图

Fig. 8 Physical device networking test environment

5.2 信息要素收集的全面性

本文对已有数据集和测试环境中拟议的攻击进行抓包分析，统计其中获取到的设备、系统和相关电力软硬件信息。统计结果如表 3 所列。主要设备为 PLC 等，本文收集了联网电力系统中可能有的相关系统和设备，初步实现了功能性较完整的验证，对各类设备进行分类以进行设备行为的学习，同时，在后续的交互中不断地扩充和完善。

表 3 收集信息统计情况

Table 3 Statistics of collected information

Information Categories	Specific type
PLC	ABB, Siemens(S7-200/300/400), Schneider(M200, M221, M340, M580, MicroLogix 1100/1400, TWDLCAA40DRF, BMX P34 2020, BMX NOE 0100, TSXETY 4103, TM221CE16R), Delta, Rockwell
Software	Industrial application software, network communication modules, HMI, SCADA Systems
Devices	Measurement and control devices, Intelligent merging units, Industrial network equipment, Industrial switches, Industrial routers, Intelligent controllers
Protocols	HTTP, Telnet, FTP, SNMP, IEC-104, Modbus, IEC-61850-GOOSE/MMS, DNP3, ICCP/TASE. 2, PROFIBUS, CIP, IEC103, IEC101, IEC92, IEC-MMS
Vulnerabilities	Weak password vulnerability, SQL injection, code execution, Trojan injection, logic vulnerability, permission acquisition

5.3 框架欺骗性评估

蜜罐实例是否足够真实决定了其能否吸引更多的攻击者来扫描探测，从而有更大概率吸引到有针对性的攻击者和执行级联和复杂攻击的攻击者。本实验将 SGPot 和部分开源蜜罐部署到互联网中进行了为期 30 天的数据采集。其中 SGPot 实例总共收集到 1678 个请求，而 GridPot 和 SHaPe 分别收集了 1379 和 955 个请求。

对于收集到的请求，本节进一步对其进行了分析，过滤掉了不符合电力行业专有协议的请求，结果如表 4 所列。

表 4 不同蜜罐实例收集请求汇总情况

Table 4 Summary of collected requests for different honeypot instances

Honeypot Instance	Number of all requests	Number of valid requests
SGPot	1678	325
GridPot	1379	273
SHaPe	955	98

为了更直观地体现请求中的各类操作，我们将所有有效请求分为两类：正常操作和威胁操作。正常操作指基本的交互和侦察请求，包括最初的探测扫描、建立连接等，这一类操作通常不会对智能电网中的各类设备和系统正常运行产生影响。威胁操作指进行的请求会影响电力设备的运行状态，所执行的操作超出了系统运行的正常范围，包括控制逻辑篡改、数据注入等。详细结果如表 5 所列，SGPot，GridPot 和 SHaPe 分别收集了 325，273 和 98 个有效的请求。可以看出，SGPot 收集到的有效请求远多于 SHaPe，而与 GridPot 相比，SGPot 收集的有效请求和威胁操作更多。

表 5 3 类蜜罐实例收集信息分类

Table 5 Classification of information collected from three types of honeypot instances

Honeypot Instances	Number of valid requests	Normal Operations		Threat Operations	
		Establishing a connection	Read status information	Control logic tampering	Data Injection
SGPot	325	151	172	0	2
GridPot	273	135	138	0	0
SHaPe	98	60	38	0	0

本文对所有收集的请求进行详细分析，按 5 天一个单位统计了请求数据，结果如图 9 所示。可以看出，统计周期内 3 个蜜罐请求数量基本保持平稳状态，说明在 30 天的时间里 3 个蜜罐都没有被识别为蜜罐。在中间一段时间请求数相对较多，其原因是在最初的侦察中，蜜罐实例通过了攻击者的侦察，并足够吸引攻击者进一步实施攻击。在最后 5 天的请求中，SGPot 收集的请求数略少于 GridPot 是因为我们对配置进行了修改，在组网环境中保证正常通信的情况下撤掉了部分设备以验证真实设备对蜜罐实例的影响。

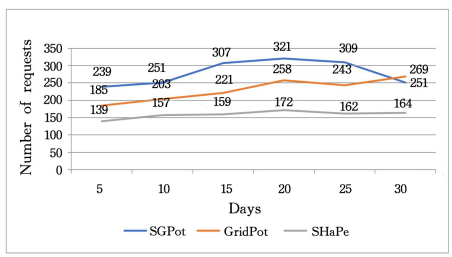

图 9 3 类蜜罐实例欺骗性对比

Fig. 9 Comparison of deceptiveness of three types of honeypot instances

5.4 会话改进

会话长度是反应攻击者在蜜罐中停留时间长短的重要判断依据，即蜜罐是否足够欺骗攻击者，而我们的目标是希望尽可能长时间地将攻击者保留在系统中，以便更长时间地研究攻击者的攻击行为。我们通过部署的蜜罐实例，分析蜜罐自身回复逻辑和经过智能决策的对比情况。

将 SGPot，GridPot 和 SHaPe 同时部署在互联网中，基于收集的信息进行统计，结果如图 10 所示。可以看出，设计的 SGPot 在会话方面有较大的提升，在会话长度小于 4 的会话中，3 个实例表现出的性能相近；在长会话中，SGPot 明显要多于其他两个；在会话长度为 6 时，SGPot 明显多于 GridPot 和 SHaPe；而在部署期间，只有 SGPot 捕获了长度大于或等于 7 的会话。

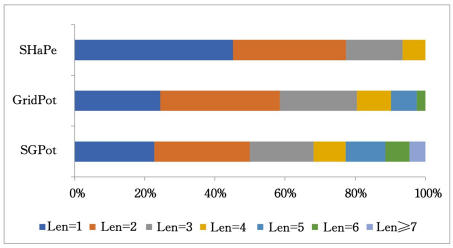

图 10 3 类蜜罐实例请求会话长度分布情况

Fig. 10 Distribution of request session length for three types of honeypot instances

5.5 业务流程仿真验证

本节的主要目的是验证 SGPot 对于智能电网业务流程仿真的有效性。实验将仿真的电网物理组件状态的实时更新情况与测试环境中的真实设备进行了比较。图 11 给出了 SGPot 和测试环境中执行相同电力操作时时间上的变化曲线。

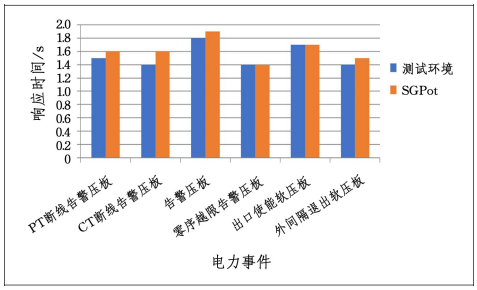

图 11 响应时间对比分析

Fig. 11 Comparative analysis of response time

从图中可以看到，SGPot 不可避免地存在一些误差，个别操作的执行时间延迟会略大于真实设备，但也是在正常操作的正常范围之内（遥控执行时间小于等于 2 s），这对于提高蜜罐实例的欺骗性来说已经足够。因为对于蜜罐来说，它并不需要完整、高精度地仿真出真实业务的全部流程，而是能够欺骗执行针对性攻击的攻击者，并诱导攻击者在初步的侦察和设备/系统识别时做出错误的判断。

6 讨论

由于电网环境的特殊性及实验条件限制，蜜罐实例的设置虽然考虑了真实环境下的攻击和防御手段，但是其只在测试环境中进行了实验验证，目前不能在真实电网环境中实际进行性能验证。我们由于在互联网上部署了蜜罐实例，因此观察到了攻击者的行为。但是，我们没有观察到诸如配置更改之类的高级攻击。然而，基于实验中显示的会话长度的比较，可以说我们的智能交互蜜罐与攻击者的交互比低交互蜜罐更有效。与高交互蜜罐不同，本文提出的蜜罐不会被破坏，也不会被攻击者劫持，它比当前用于观察攻击的高交互方法更安全。

结束语 智能电网中的攻击者在执行针对性攻击时，需要在发起攻击之前对电网的拓扑及其中的设备信息进行侦察。如果没有与攻击者较好的交互机制，就很难捕获完整的漏洞攻击代码和攻击链。基于此，本文提出了一种利用数据集及蜜罐实例收集潜在响应的方法，并利用强化学习的方法在与攻击者交互的过程中学习正确和预期的行为。实验表明，所提框架可以改善与攻击者的会话并捕获更多的攻击。

我们的目标是实现一个智能电网的通用蜜罐框架，并结合进一步改进的测试环境来提高框架的工作效率，进一步完善并在真实运行的电力环境中进一步验证。由于部署时间有限，我们获取到的数据量不大，后期的研究将基于探测策略迭代的方式来优化信息的收集，同时重点对收集的信息进行更好的处理。

参 考 文 献

[1] CASE D U. Analysis of the cyber attack on the Ukrainian power grid[J]. Electricity Information Sharing and Analysis Center (E-ISAC),2016,388:1-29.

[2] AN T. Comprehensive Analysis Report on Attacks on Ukraine's Power System [R]. 2016,2016.

[3] KAZI R,KUMAR N. Thinking the Unthinkable:Cyber Attacks on India's Nuclear Assets[J]. Liberal Stud. ,2019,4:107.

[4] LI F,YAN X,XIE Y,et al. A review of cyber-attack methods in cyber-physical power system[C]// 2019 IEEE 8th International Conference on Advanced Power System Automation and Protection(APAP). IEEE,2019:1335-1339.

[5] PIETROSEMOLI L,RODRÍGUEZ-MONROY C. The Venezuelan energy crisis:Renewable energies in the transition towards sustainability[J]. Renewable and Sustainable Energy Reviews, 2019,105:415-426.

[6] BUZA D I,JUHÁSZ F,MIRU G,et al. CryPLH:Protecting smart energy systems from targeted attacks with a PLC honeypot[C] // International Workshop on Smart Grid Security. Cham:Springer,2014:181-192.

[7] KOŁTYŚ K,GAJEWSKI R. Shape:A honeypot for electric power substation[J]. Journal of Telecommunications and Information Technology,2015(4):37-43.

[8] REDWOOD O,LAWRENCE J,BURMESTER M. A symbolic honeynet framework for scada system threat intelligence[C]// International Conference on Critical Infrastructure Protection. Cham:Springer,2015:103-118.

[9] MASHIMA D,CHEN B,GUNATHILAKA P,et al. Towards a grid-wide,high-fidelity electrical substation honeynet[C]// 2017 IEEE International Conference on Smart Grid Communications

(SmartGridComm). IEEE, 2017: 89-95.
[10] MASHIMA D, LI Y, CHEN B. Who's Scanning Our Smart Grid? Empirical Study on Honeypot Data[C]// 2019 IEEE Global Communications Conference (GLOBECOM). IEEE, 2019: 1-6.
[11] MASHIMA D, KOK D, LIN W, et al. On design and enhancement of smart grid honeypot system for practical collection of threat intelligence[C]// 13th USENIX Workshop on Cyber Security Experimentation and Test(CSET 20). 2020.
[12] GUNATHILAKA P, MASHIMA D, CHEN B. Softgrid: A software-based smart grid testbed for evaluating substation cybersecurity solutions[C]// Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy. 2016: 113-124.
[13] LIN H, ZHUANG J, HU Y C, et al. DefRec: Establishing Physical Function Virtualization to Disrupt Reconnaissance of Power Grids' Cyber-Physical Infrastructures[C]// Proceedings of 2020 Network and Distributed System Security Symposium(NDSS). 2020.
[14] TC57 I E C. IEC 61850-90-2 TR: Communication networks and systems for power utility automation-part 90-2: Using iec 61850 for the communication between substations and control centres [S]. International Electro technical Commission Std, 2015.
[15] "CONPOT ICS/SCADA honeypot," [EB/OL]. https://www.conpot.org.
[16] WAGENER G. Self-adaptive honeypots coercing and assessing attacker behaviour[D]. Institut National Polytechnique de Lorraine-INPL, 2011.
[17] LUO T, XU Z, JIN X, et al. Iotcandyjar: Towards an intelligent-interaction honeypot for iot devices[J]. Black Hat, 2017, 1: 1-11.
[18] PAUNA A, BICA I. RASSH-Reinforced adaptive SSH honeypot [C]// 2014 10th International Conference on Communications (COMM). IEEE, 2014: 1-6.
[19] PAUNA A, IACOB A C, BICA I. Qrassh-a self-adaptive ssh honeypot driven by q-learning[C]// 2018 International Conference on Communications(COMM). IEEE, 2018: 441-446.
[20] PAUNA A, BICA I, POP F, et al. On the rewards of self-adaptive IoT honeypots[J]. Annals of Telecommunications, 2019, 74(7): 501-515.
[21] YAMAMOTO M, KAKEI S, SAITO S. FirmPot: A Framework for Intelligent-Interaction Honeypots Using Firmware of IoT Devices[C]// 2021 Ninth International Symposium on Computing and Networking Workshops (CANDARW). IEEE, 2021: 405-411.
[22] ANTONIOLI D, TIPPENHAUER N O. MiniCPS: A toolkit for security research on CPS networks[C]// Proceedings of the First ACM Workshop on Cyber-physical Systems-security and/or Privacy. 2015: 91-100.
[23] KAUR K, SINGH J, GHUMMAN N S. Mininet as software defined networking testing platform[C]// International Conference on Communication, Computing & Systems(ICCCS). 2014: 139-142.

WANG Yuzhen, born in 1998, postgraduate. His main research interests include smart grid security and deception defense technologies.

WEI Qiang, born in 1979, Ph.D, professor, doctoral supervisor. His main research interests include software vulnerability analysis and vulnerability mining, industrial Internet security, etc.

(责任编辑:何杨)